Dihoru, L., Dietz, M., Horseman, T., Kloukinas, P., Oddbjornsson, O., Voyagaki, E., Crewe, A. J., & Taylor, C. A. (2018). Neural networks for displacement analysis in an advanced gas cooled reactor core model. *Nuclear Engineering and Design*, *332*, 252-266. https://doi.org/10.1016/j.nucengdes.2018.03.039

Neural Networks for Displacement Analysis in an Advanced Gas Cooled Reactor Core Model

Luiza Dihoru*, Matt Dietz, Tony Horseman, Panos Kloukinas, Olafur Oddbjornsson, Elia Voyagaki, Adam J. Crewe, Colin A. Taylor

University of Bristol, University Walk, Bristol, BS8 1TR, United Kingdom
* Corresponding author: Luiza.Dihoru@bristol.ac.uk

Abstract
This paper presents a Neural Network (NN) approach for displacement analysis with applications in modelling the seismic response of the UK's Advanced Gas Cooled Reactors (AGRs). A quarter sized physical model of a reactor core was developed at the University of Bristol to provide experimental validation to the existing numerical models that support the seismic resilience assessments of the AGRs. The physical model outputs include displacement and acceleration datasets of considerable size and complexity, collected for a range of seismic inputs and postulated component damage scenarios. Rich sets of displacement data were employed in training two NN models that can predict displacement at user-defined locations in the core physical model and can map the correlation between the component relative displacements. Understanding component displacements is particularly important, as such displacements may affect the channel shapes and can cause local and general distortion of the core. This paper presents the development, testing and performance of the NN models. The NNs yield predictions that compare well with the experimentally obtained parameters. As more experimental test data become available, the NN's prediction capability will benefit from accumulated training. In the future, the NNs will be incorporated into a multi-layered framework for dynamic response prediction and analysis.

Keywords: advanced gas cooled reactor, seismic testing, neural network,displacement prediction

1 Introduction

The aim of this paper is to test the validity of use of NNs as predictor tools for a complex multi-rigid body dynamics problem. Neural networks (NNs) are widely used in all areas of engineering and experimental research to model systems with numerous variables and complex non-linear response, and their role could be of pattern recognition, object classification or prediction (Beale and Jackson 1990). NNs act as nonlinear modelling systems that consist of processing units or neurons interconnected by means of weights. The model is represented by the values of these weights, which are established during network training. How the strengths of the neuron connections are obtained during the training phase to achieve a desired overall behavior of the network is governed by the training algorithm (Beale and Jackson 1990, Beale *et al* 2016). Each neuron has an associated transfer function that relates its inputs to its output. Given a set of inputs, a NN can predict the outputs without the need for a priori assumptions with regards to the relationship between the inputs and the outputs. The different types of NN architecture, transfer functions and training algorithms in use today form a vast area of mathematical research (see, for example, Beale and Jackson 1990, Heaton 2012 and Hagan et al 2014).

In the nuclear industry, a plethora of computer systems based on object classification NNs have been proposed as diagnostic tools for the direct mapping of plant signals into components faults (see for example Uhrig 1991a, Uhrig 1991b, Bourguet and Antsaklis 1994, Santosh *et al* 2007). NNs have also been applied for the characterization of normal operating plant conditions and transient data pre-processing, and have been used in combination with other computer tools in hierarchical multi-level diagnostic systems (Cheon and Chang 1993, Ohga and Seki 1993, Bartal *et al* 1995, Santosh *et al* 2007). Several NN-based systems have been proposed to deal with off-line nuclear power plant - related issues, such as fuel reload optimization, power and heat distribution in the fuel elements, fault

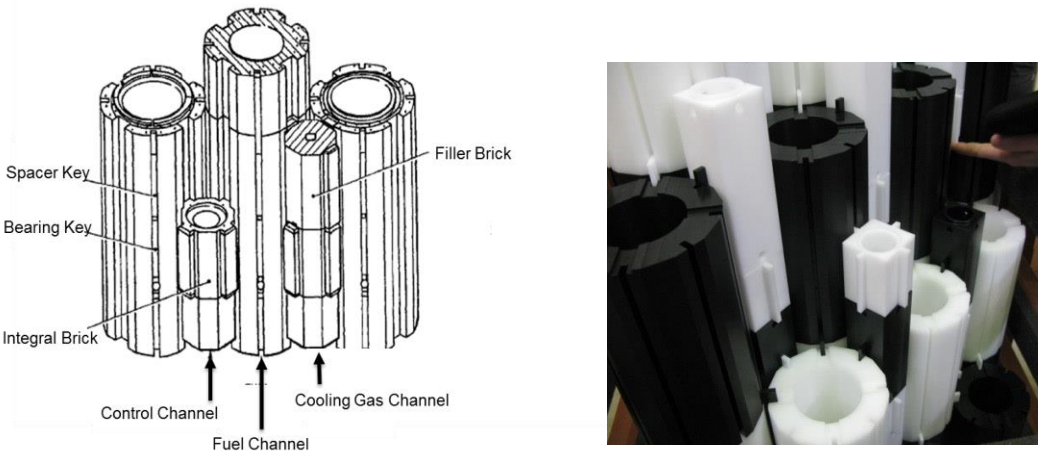
detection in fuel elements or steam header welds (Reifman *et al*,1994, Reifman *et al* 1996). Other more recent projects discuss the feasibility of using neural networks for modelling the plant thermodynamics and for the analysis of plant vibrations (Westwick D (ed) (2007)). In the UK, a novel approach to generate models of systems for real time monitoring in the nuclear industry has been developed using tools based on binary neural networks and associative memory techniques (Cybula 2016).The most widely used type of neural network, both for nuclear plant diagnostics and plant model identification, is the feedforward back propagation NN which consists of a layered arrangement of neurons. The information flows unidirectionally from the input layer of neurons to the output layer, while the error between the NN prediction and the values of the training set propagates backward through the network. This type of NN has been classified as an universal approximator (Hornik *et al* 1989).
This paper presents the use of the NN approach in the context of physical modelling of an Advanced Gas Cooled Reactor (AGR) core. The AGRs are the second generation of British gas-cooled nuclear reactors, using graphite as the neutron moderator and carbon dioxide as the coolant. In the United Kingdom, there are seven AGR power stations, each with two identical reactors. Their cores comprise of a stacked array of hollow, prismatic graphite bricks and interlocking keys and are designed to fulfil three fundamental functions: provide neutron moderation, allow movement of fuel and control rods and allow gas flow for cooling. These functions must be maintained in normal operating conditions, but also during hazardous natural events, such as earthquakes. According to current international standards, a nuclear plant should be qualified against at least 0.1g peak ground acceleration, while the plant operators require that their AGRs can be safely shut down and held down in the case of a more severe seismic event with a probability of exceedance of 10^{-4} per annum.The seismic capability of the AGRs is subject to continual examination by the operators as part of a strategy that involves the development of enhanced analytical methods and physical models for static and dynamic behaviour (Neighbour 2007, Neighbour 2013, Flewitt and Wickham 2015). The earliest physical model for seismic behaviour included a simple 9x9 brick array, employed by the National Nuclear Corporation (NNC) in 1985 (Rogers 2012). Later on between 2008-2012 a single layer 20-rings array model and a 4x4x8 array ('the minicore') were developed by the University of Bristol (UOB) to explore the basic mechanics of the core system and prepare the ground for a more sophisticated modelling tool (Dihoru *et al* 2016).

In 2014, a quarter scale 8-layer-20-bricks-across AGR core model ("the Multi Layer Array (MLA) rig") was commissioned at the UOB for the investigation of the seismic behaviour of these complex systems. The rig can generate brick displacements of sufficient magnitude to exceed current seismic assessment limits when simulating the effects of postulated component degradation (i.e. cracked bricks and/or failed keys) and increased brick-to-brick clearances arising from radiation shrinkage in the AGR cores. The MLA was tested on the earthquake simulator ('shaking table') at UOB. It provided useful insights into the dynamics of core arrays and feeds large amount of data into the existing numerical models (Kralj et al 2005, Koziara and Bićanić 2011) for validation.

Nine MLA experimental configurations comprising ~42000 components and ~3200 measurement sensors generated ~7Tb of data. The physical model outputs include displacement and acceleration datasets of considerable size and complexity that feature all the challenges associated with 'big data' capture, storage, transfer, analysis, visualization, updating and information privacy. The vast amount of data that resulted from the tests are challenging in both handling and interpretation and require bespoke software for structuring and analysis. Besides the more conventional processing tools, a new neural network (NN) approach has been developed to predict the displacement of bricks in the top layer of the array, based on past response history.Such predictions are important in understanding the expected reponse for a user-defined seimic input and for recognizing the correlations between the individual component displacements at various locations in the array. Such displacements may reflect the local and the general core distortion, as well as being an indicator of the behaviour of the top of the control and fuel channels.The NN models and the details of NN training and testing are presented together with the NN simulation results. The NN's prediction performance shows that the new models are efficient in dealing with highly non-linear, time-dependant displacement data and that they are capable of recognizing patterns of behaviour determined by both geometry and dynamic input.

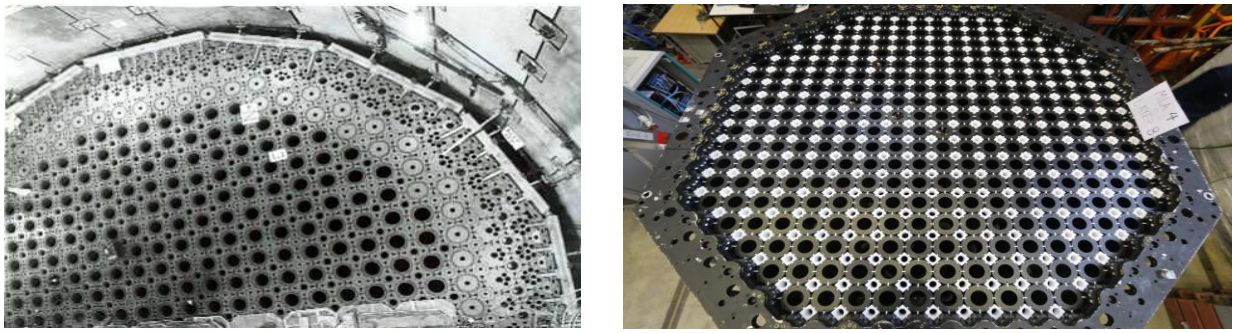
2 MLA Experimental Rig

The MLA consists of an 8-layer assembly of quarter scale model bricks and model keys made of a rigid engineering plastic (acetal). All of the AGR graphite component types are represented in the MLA, i.e. fuel (lattice) bricks, interstitial bricks, filler bricks, filler, spacer and loose bearing keys (Figure 1). The multi-layered array of components is octagonal in shape and has 20 bricks across its maximum cardinal direction, modelling the innermost 10 rings of the AGR core (Figure 2). The array is enclosed by a rigid support frame and confined at the bottom by a rigid arrangement of plastic plates that permit the required radial rocking and in which the bottom component of each vertical column is mounted (Figure 3). More details on the MLA's design and mode of operation can be found in Dihoru *et al* 2011, Dihoru *et al* 2012, Dihoru *et al* 2014, Dihoru *et al* 2016.



a) b)

Figure 1 Arrangement of AGR core components (a) and MLA model components made of acetal (b)

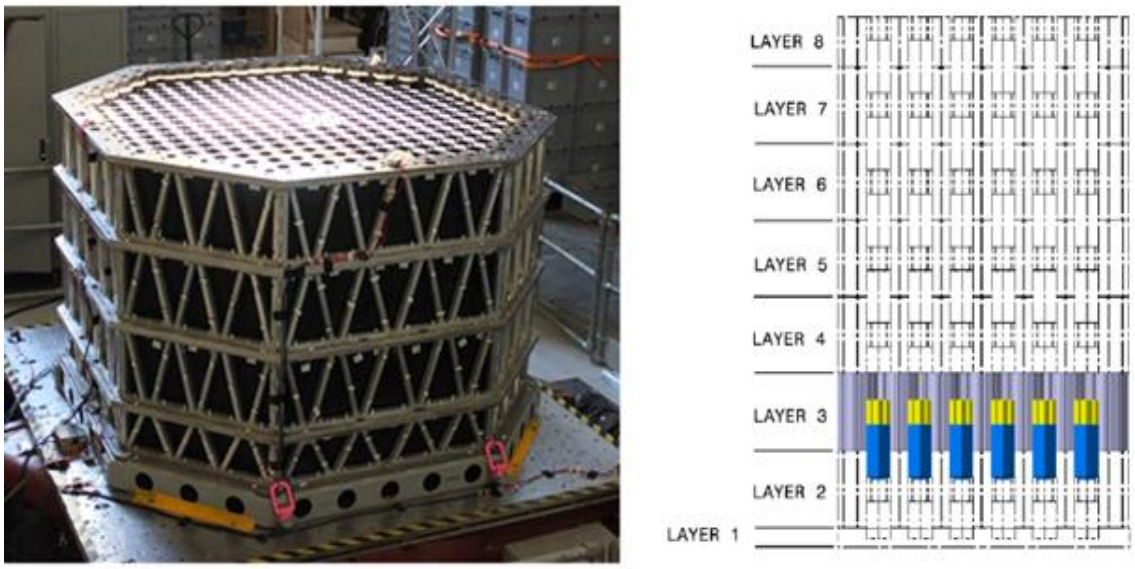


a) b)

Figure 2 An AGR core (a), courtesy EDF Energy), the MLA rig (b), view from above.

For dynamic testing, the MLA rig is secured by its base to the platform of an earthquake simulator (Figure 3a) and subjected to simulated seismic motions of various frequency and acceleration ranges applied in different directions in the horizontal plane. During testing, accelerations and displacements are recorded at relevant locations inside the array and on the top layer of the array by a variety of instruments (accelerometers, Hall effect sensors, infrared markers and high speed cameras). During a seismic event, the MLA behaves as an array of rigid bodies in which the relevant forces are the impact forces generated during the collisions between the components, the gravitational and the restoring forces. The energy loss after a brick-to-brick collision depends heavily on the actual layout of components in a zone of investigation (i.e. presence or absence of keys, locking of keys, etc). In

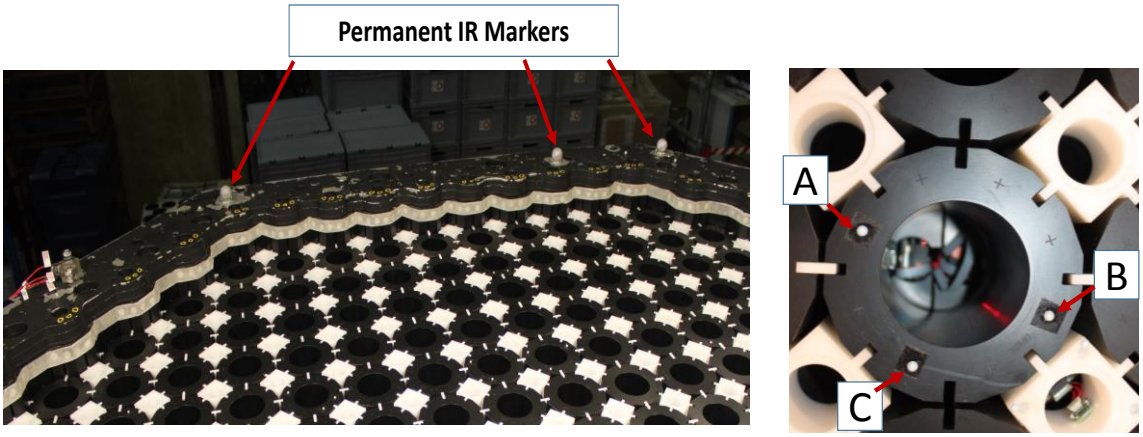
principle, the MLA's measurement systems target the local and the general distortion of the core, as well as the control and lattice channel shapes.



a) b)

Figure 3 The MLA's rigid restraint frame (a). Layer layout in the MLA rig (b). Layer 1 is an assembly of plastic plates. Layers 2-8 are active. Layer 3 is highlighted to show arrangement of bricks. (courtesy Atkins).

The top layer of the MLA is monitored via a 5-camera vision system that records the motion of the infrared (IR) passive markers attached to selected lattice bricks and to the MLA's restraint frame. (Figure 4). The infrared cameras work at a speed of 100 frames/second in conjunction with a tracking software that generates output files containing the X, Y, and Z coordinates of each IR marker. Each brick that is monitored contains three IR markers, hence its 3D position can be fully determined.



a) b)

Figure 4 Infrared (IR) markers mounted on the MLA restraint frame (a), IR markers (A, B and C) mounted on the model fuel brick (b).

The displacement measurement system for the MLA's top layer has generated a vast amount of data stemming from testing nine array configurations in test schedules containing ~400 tests each. The data reflect the non-linear and noise-contaminated response of the MLA. Data sets such as these are

typically well-suited for a NN modelling approach. The authors have pursued this approach for predicting displacement in the top layer for a user-given seismic input. The details of the new NN tools are given in the following sections.

3 Neural Networks for MLA Displacement Analysis

3.1 Work Objectives

The NN approach presented in this paper deals with an analysis of component displacement in the top layer of the MLA for user-defined seismic inputs. Displacement of top layer components is an important model output that gives an indication of the local and general distortion of the core and also on the deformation at the top of the fuel and control channels. In the AGR cores, the channel shapes have to be kept within tightly controlled limits for the safe insertion of the control rods and the fuel stringers. It is important to explore the core displacements in the quarter scaled model (MLA) and translate the results to the prototype numerical models, for different array configurations and for arrays with different amounts of postulated damage (i.e. different distributions and levels of brick cracking). The top layer reponse may be a good indicator of the way the effect of damage propagates inside the array. In general, displacement prediction is not a trivial exercise: the dynamic behaviour of the MLA is the result of a highly complex set of factors of influence that are summarized below.

a) ***Intricate pathway of dynamic force transmission:*** the seismic input is a displacement time history applied by the earthquake simulator at the base of the rig frame. The MLA array consists of vertical columns rigidly fixed at the base and enclosed by the rigid rings of the frame. The keys connecting the bricks in the horizontal plane serve as a radial system of component-frame interaction. When the base seismic input is applied, a primary dynamic input is transmitted up the columns (which act as flexible beams) causing the brick to move horizontally thus engaging the radial keying system. When the radial keying system is engaged, a secondary force path is developed in the horizontal plane with the dynamic input being transmitted from the lateral wall of the frame to the array. The most important forces inside the array are the collision forces, the inertial forces and the gravitational forces. The collisions are complicated due to their 'spiky' nature (high frequency content). Also, of a lesser significance for the overall response, but still important for the local response, are the friction forces at the key-to-brick and brick-to-brick interfaces. These interactions are complicated to model analytically due to their non-linear and transient nature, the large number of degrees of freedom and the large number of component-to-component interfaces.
b) ***Complex component geometrical layout***: the MLA array consists of ~44000 components (bricks and keys) stacked vertically in columns and arranged in radially- distributed rings. The clearance between the vertical fuel columns is constant (4mm). Due to the gap accumulation effect, the components situated in the central area of the core are expected to move more that the ones in the vicinity of the rigid restraint frame. The model fuel bricks have end face features ('rocking features') that cause them to rock preferentially in the array radial direction. Depending on the location of the brick in the horizontal plane, the 'rocking feature' direction changes and this contributes to components of motion that are not necessarily aligned with the seismic motion. An individual brick will tend to move on the direction of motion (as part of a column), but will translate and rotate relative to the brick belowin accordance with its end face feature. This combination of motions makes the relative displacements of a component difficult to predict using simple structural dynamic models.

Given the complexity of the MLA's dynamic response and the presence of a large amount of experimental data, a NN approach for modelling displacement was considered particularly suitable, as such an approach does not require an *a priori* assumption regarding the relationship between the input and the output variables. The proposed NN models are presented in Table 1.

The architecture details for the two NN models are presented in Section 4.2. In the future, the NNs can be incorporated into a multi-layer framework for dynamic system response that can fulfil four distinct roles:

a) A pattern recognition tool that identifies patterns and irregularities in response;
b) A prediction tool that evaluates the response for previously unseen seismic inputs;

c) An experimental design tool that identifies relevant inputs to be employed in future experiments;
d) A 'big data' investigation tool that identifies the most interesting tests for future data analysis, based on pre-defined selection criteria.

This paper deals with the development of the pattern recognition and the prediction tools (points a) and b)), while the experimental design tool and the 'big data' investigation tool will be covered in future publications. The 'big data' investigation framework will build on the historic accumulation of knowledge facilitated by the previous tools and will include the following strands of work: data capture, data storage, data analysis, data sharing and transfer, data visualization and data querying. In particular, the data analysis will benefit from navigation and interrogation tools meant to reveal certain patterns and trends based on pre-defined selection criteria. The 'big data' framework should be general and complex enough to be applicable to a wider set of AGR-related issues.

Table 1 Proposed neural networks for displacement analysis

ID	Proposed Models	Inputs	Output
NN1	NN for displacement prediction in top layer	Applied seismic input and brick location	Absolute displacement of bricks at selected locations in the top layer
NN1 Objective: By cumulative training, the NN1 model can predict the response for a large number of seismic waves, substituting or complementing the experimental testing. If predictions show interesting features of response, then certain selected inputs can be used in physical experiments to measure response.			
NN2	NN for correlation maps	Displacement relative to the frame for the brick in central position	Displacement relative to the frame for bricks at selected locations in the top layer. Regression coefficients measuring correlation of relative displacement between brick in central position and other bricks at selected locations in the top layer.
NN2 Objective: Prediction of correlation between relative displacements of components in the top layer. This helps with understanding how the seismic input and/or the cracking in the lower layers affect the dynamic response in the top layer.			

3.2 Work Planning

One intact array and four cracked array configurations were selected for this investigation (they are identified by different IDs in Table 2). These configurations were employed in order to test the validity of the NN method for different study cases and also to seek features of response that may be specific for a certain type of array distribution. The chosen cracked arrays model postulated damage scenarios of an aged core containing doubly cracked bricks or singly cracked bricks. The cracked configurations were selected on purpose, as their displacement response is larger and has more variety across the array than an intact configuration. The pattern of cracking in MLA2, MLA3 and MLA4 was computer generated, based on a random number distribution algorithm. Thus, the physical arrays MLA2, MLA3 and MLA4 contain layouts of cracking with an approximately uniform spatial distribution in layers 4-7. The MLA8 configuration included singly cracked bricks in a cluster situated in the north part of the array, while the rest of the array was intact.

The NN training employed absolute and relative displacements recorded experimentally at 19 locations in the array, for the model fuel bricks B1-B19 (Figure 5). The selection of the 19 locations was done so that each quadrant of the array is representatively covered in the analysis. Both the central and the boundary areas of the array are represented.

Table 2 Array configurations under investigation

No.	ID	Top Layer (Layer 8)	Damage Percentage in the Array (Layers 4-7)	Damage Distribution
1	MLA1	intact	No damage (uncracked array)	
2	MLA2	intact	30 % doubly cracked bricks	Random (approx. uniform)
3	MLA3	intact	30 % doubly cracked bricks	Random(approx. uniform)
4	MLA4	intact	50 % doubly cracked bricks	Random(approx. uniform)
5	MLA8	intact	5 % singly cracked bricks arranged in a cluster on the north side of the array	Uneven

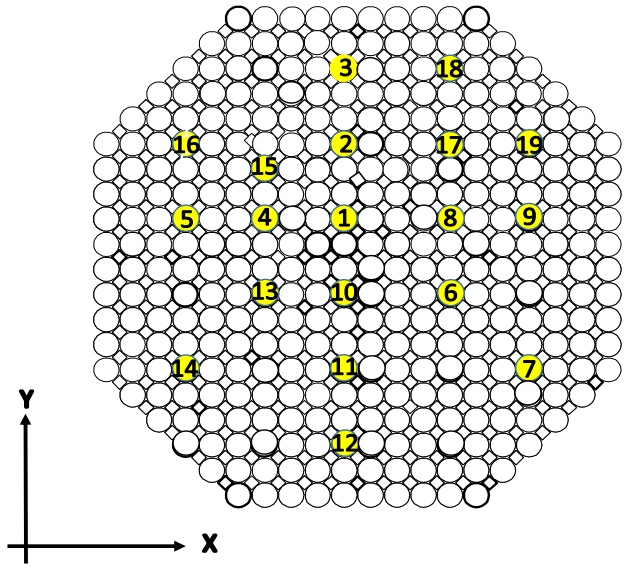
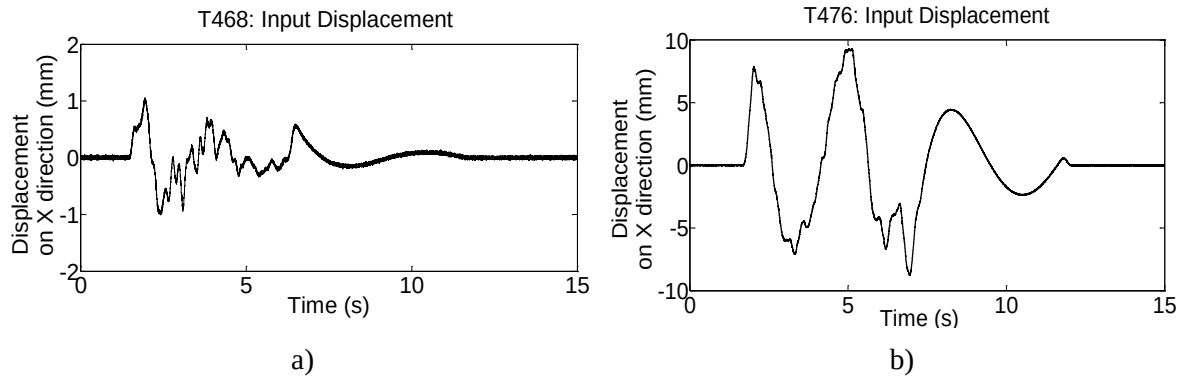


Figure 5 Mapping of model bricks in the top layer of the MLA array. The displacement records of Bricks 1-19 (as marked) were employed in training the neural networks.

The NN assimilates data from bricks located in both central and border regions to encapsulate a wider displacement response. The absolute displacements are the brick displacements in the fixed infrared camera system of coordinates, while the relative dispalcements represent the brick displacements relative to the rig frame.

The rig frame is rigidly attached to the shaking table platform. The rigid construction of its base plate and outer rings means that the input of the shaking table is transmitted fully to the model array via the frame. In the following analysis, the seismic inputs applied to the array are collected via the infrared markers rigidly attached to the top of the frame. A number of seismic inputs were employed in the investigation which were derived from the hazard inputs and resulting responses predicted by the seismic assessments of the UK's AGR stations (Principia Mechanica Limited 1981). They were in the form of time histories generated from secondary response spectra at the core boundary for seismic events with a 10^{-4} probability of occurrence (labelled 'HPB'). In addition to those, a set of Eurocode required response spectra (RRS) compatible time histories (labelled 'RRS') were also employed (see Figure 6). All inputs were applied by the earthquake simulator on a single axis (X or Y direction). For each configuration presented in Table 2, data from 10 tests on X direction and data from 10 tests on Y direction were employed in the analysis. Various amplitude amplification factors were applied to the input time histories. The maximum input displacement in tests varied beween 1-30 mm.



a) b)

Figure 6 Examples of displacement inputs as measured by infrared makers on the rig frame. Inputs employed in MLA2: a): HPB, b): RRS compatible.

3.3 NN Architecture, Mode of Operation and Results

The neural network codes were written in Matlab™ , R2012a, Version 7.14, and the NN training and simulation used the Neural Network Toolbox associated with the same software package.

3.3.1 NN1 – Model for Displacement Prediction

Figure 7 presents the NN's mode of operation. The inputs of NN1 are the input displacement time history at the top of the frame and the locations of seven selected bricks (B1-B7) in the top layer. NN1 presents 7 outputs which are the absolute displacements of bricks B1-B7. The displacement time histories employed in the analysis are experimental data measured by the IR vision system at a speed of 100 frames per second. Hence, the time step of the data vectors is 0.01s. The NN architecture was empirically selected to include 3 hidden layers of 8 neurons each, an output layer of 7 neurons and 'weights' that are multipliers and 'biases' which are scalars added to the input signal to displace the curve of the neuron's transfer function (Figure 8). The chosen architecture is a standard one and is not dependent on the variable characteristics. According to some authors, three layers of active units can represent any pattern classification (Beale and Jackson 1990). The NN's mode of operation is based on the random spliting of the experimental data in two equal sets: the training set and the test set. The training set is employed for learning the pattern connecting the inputs to the outputs, while the test set is used in a post-training analysis (regression analysis) to evaluate the network's performance. A typical sigmoid transfer function was employed in the internal layers (Equation 1). This has the advantage of compressing a wide range of inputs into a finite range by transforming inputs that may have any value between plus and minus infinity into ones between 0 and 1. This transfer function is commonly used in the backpropagation networks, in part because it is differentiable.

$$f(x)=(1+e^{-\lambda x})^{-1} \quad (1)$$

where λ is the gain coefficient employed to control the slope of the transfer function.

The training session took place in four stages: assembling the training data, creating the network object, training the network and simulating the NN response to new inputs. The NN uses the default performance function associated with the feedforward networks [Beale et al 2016]. The performance function is the mean square error between the outputs of the network and the target outputs from the experimental data set, defined by:

$$\boldsymbol{E_k} = 0.5 \sum_i (a_i - t_i)^2 \quad (2)$$

where $\boldsymbol{E_k}$ is the global network error after $\boldsymbol{k}$ training iterations (epochs of training), $\boldsymbol{a_i}$ is the NN prediction of output i and $\boldsymbol{t_i}$ is the target value of output $\boldsymbol{i}$.

The NN training employed the standard backpropagation with the gradient descent algorithm. This is a commonly used algorithm that involve passing the errors back down the network and adjusting the weights between the neurons in the direction of the negative gradient of the performance function at each iteration (Equations 3 &4). At the beginning of training, the weights and biases are initialized with random values. Then the set of training inputs is shown to the network. During each training iteration (epoch), the network uses its transfer functions to approximate the relationship between the given inputs and the target outputs. At the end of each epoch the error between the NN outputs and the target outputs times the learning rate is used to adjust the weights of the network. The learning rate is a multiplier with values between 0 and 1 that determines the rate at which the NN converges towards a stable solution.

$$w^{[i]}{}_{mn}{}^{[k]} = w^{[i]}{}_{mn}{}^{[k-1]} - \eta\, g^{[i]}{}_{mn} \quad (3)$$

$$g^{[i]}{}_{mn} = \partial E_k / \partial\, w^{[i]}{}_{mn}{}^{[k-1]} \quad (4)$$

where $\boldsymbol{w^{[i]}{}_{mn}{}^{[k]}}$ is the weight connecting neuron no. ***m*** from ***(i-1)th*** layer to neuron no. ***n*** from ***ith*** layer at the end of epoch no. ***k***, $\boldsymbol{g^{[i]}{}_{mn}}$ is the current gradient of the error surface, $\boldsymbol{\eta}$ is the learning rate and $\boldsymbol{E_k}$ is the global error between the NN output and the target output after ***k*** learning sessions.

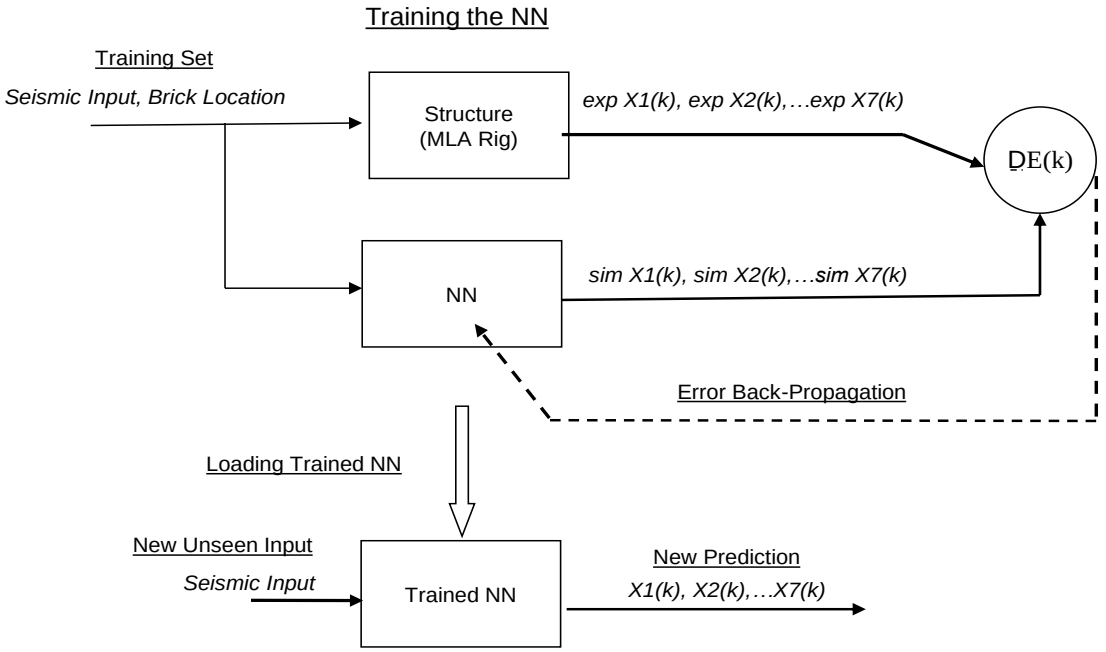


Figure 7 Dataflow diagram showing the rig data ('*exp*') for training the NN, the NN simulated data (*'sim'*) and the propagation of error back to the NN. The trained NN is then employed for a new prediction.

The algorithm of gradient descent was implemented in batch mode, i.e. all the inputs were applied to the NN before the weights were updated. Training times varying between 5000 and 50000 epochs were used.

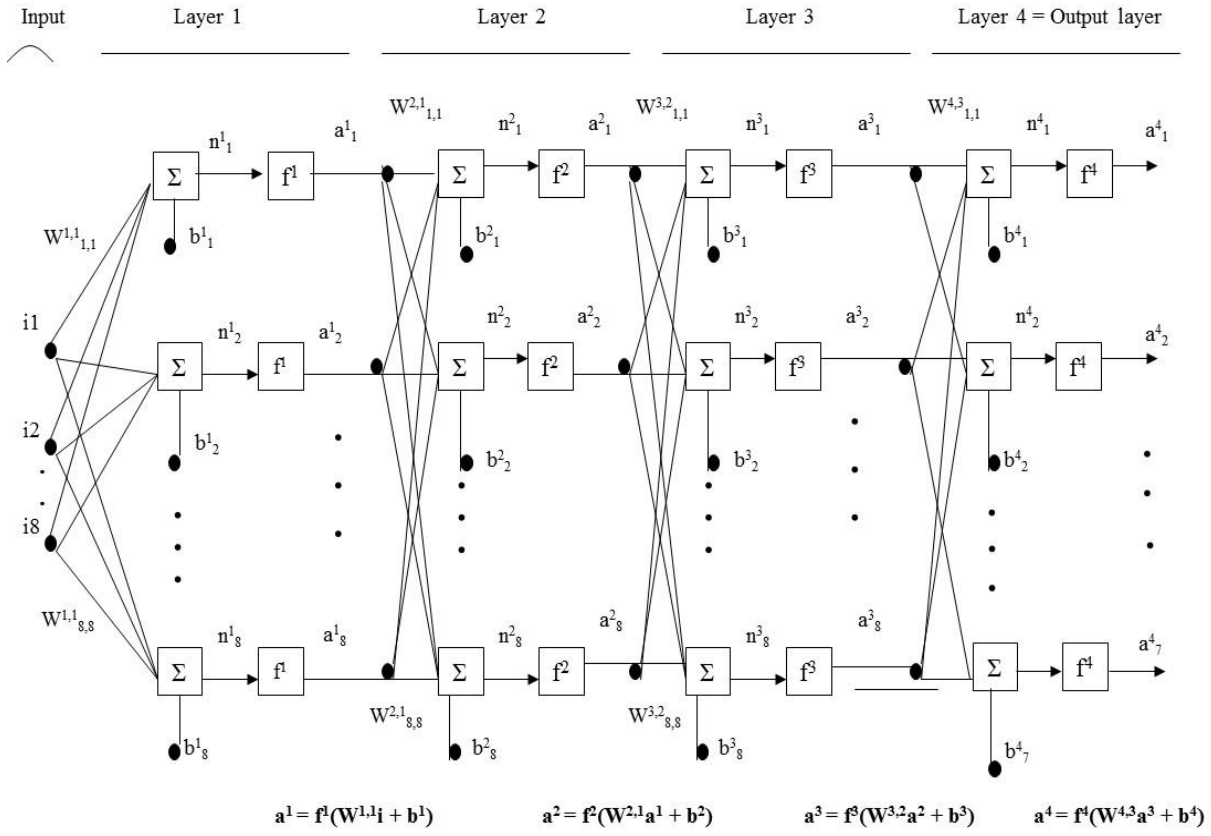
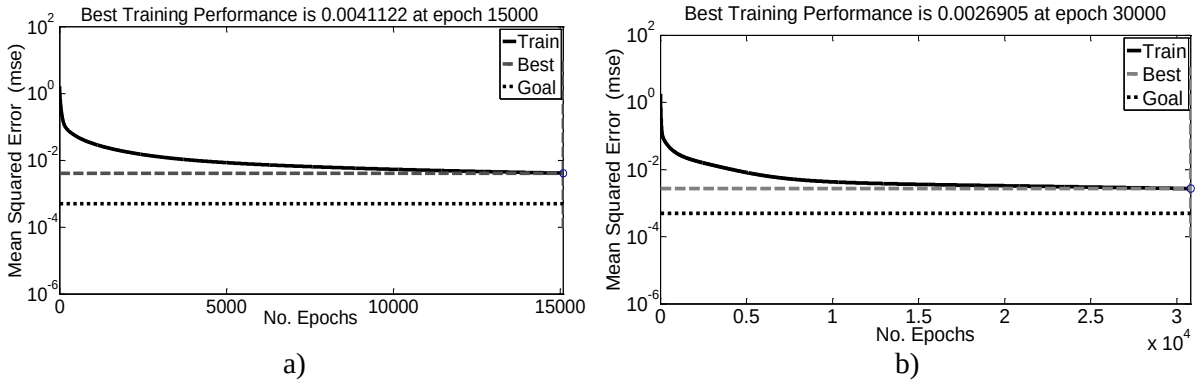


Figure 8 Architecture of neural network (NN1) with 8 inputs ('*i*'), 7 outputs ('*a*'), 3 internal layers of neurons and 1 output layer of neurons.

Figure 9 presents the evolution of the network's performance function at two times during training. The mean square error between the NN's predicted outputs and the target outputs decreases with the number of training iterations.



a) b)

Figure 9 Evolution of NN performance during training after 15000 epochs (a) and after 30000 epochs (b).

Figure 10 presents the NN's performance after 50000epochs with no failed validation checks.

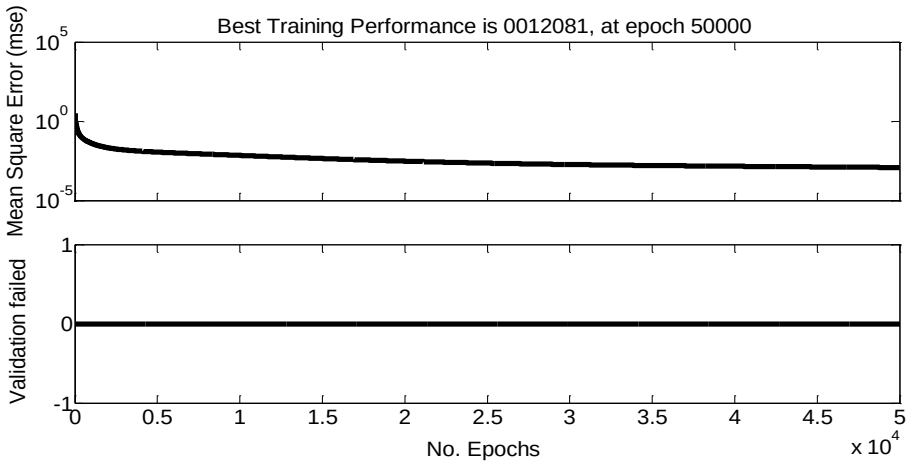


Figure 10 NN performance after 50000 epochs (top) and plot of failed validation checks (bottom).

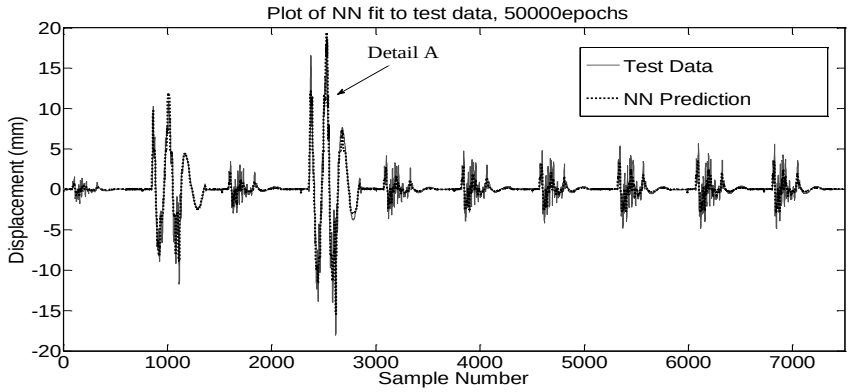
Each simulation was followed by a regression analysis to compare the simulated displacements with the experimental values belonging to the test set. The regression analysis returned the following equation relating the NN simulated data (*O*=Output) to the test set of experimental data (*T*=Target):

$$O = mT + b \tag{5}$$

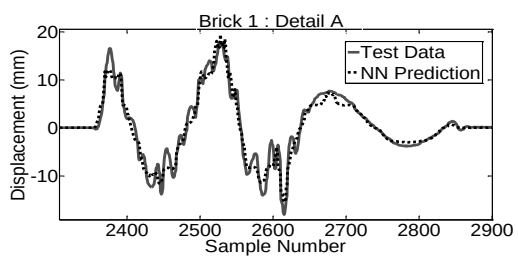
where **m** and **b** correspond to the slope and the y-intercept of the best linear regression relating targets to the network outputs. The slope of the fitted line is equal to the correlation between the Output and the Target corrected by the ratio of standard deviations of these variables. The regression coefficient (***R***) was chosen as a measure of NN performance (Equation 6).

$$R = covariance\ (Output, Target) /\ [SD\ (Output) \times SD\ (Target)] \tag{6}$$

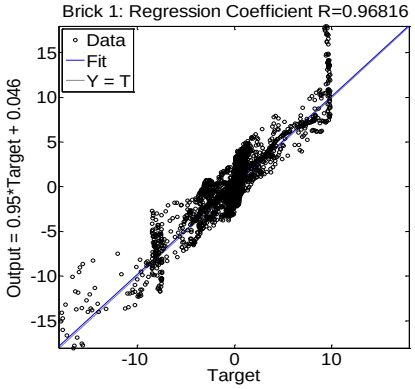
where ***SD*** is the standard deviation of the mentioned variable.The NN prediction results for Brick no.1 can be found in Figure 11.



a)



b)



c)

Figure 11 Experimental data (test data) and NN simulated data for Brick 1 displacement for 10 seismic tests, after 50000 epochs of training (a and b). Regression analysis between NN prediction (*O*=Output) and experimentally masured target data (*T*=Target) for Brick 1(c).

The regression analysis results for Bricks 2-7 are shown in Figure 12.

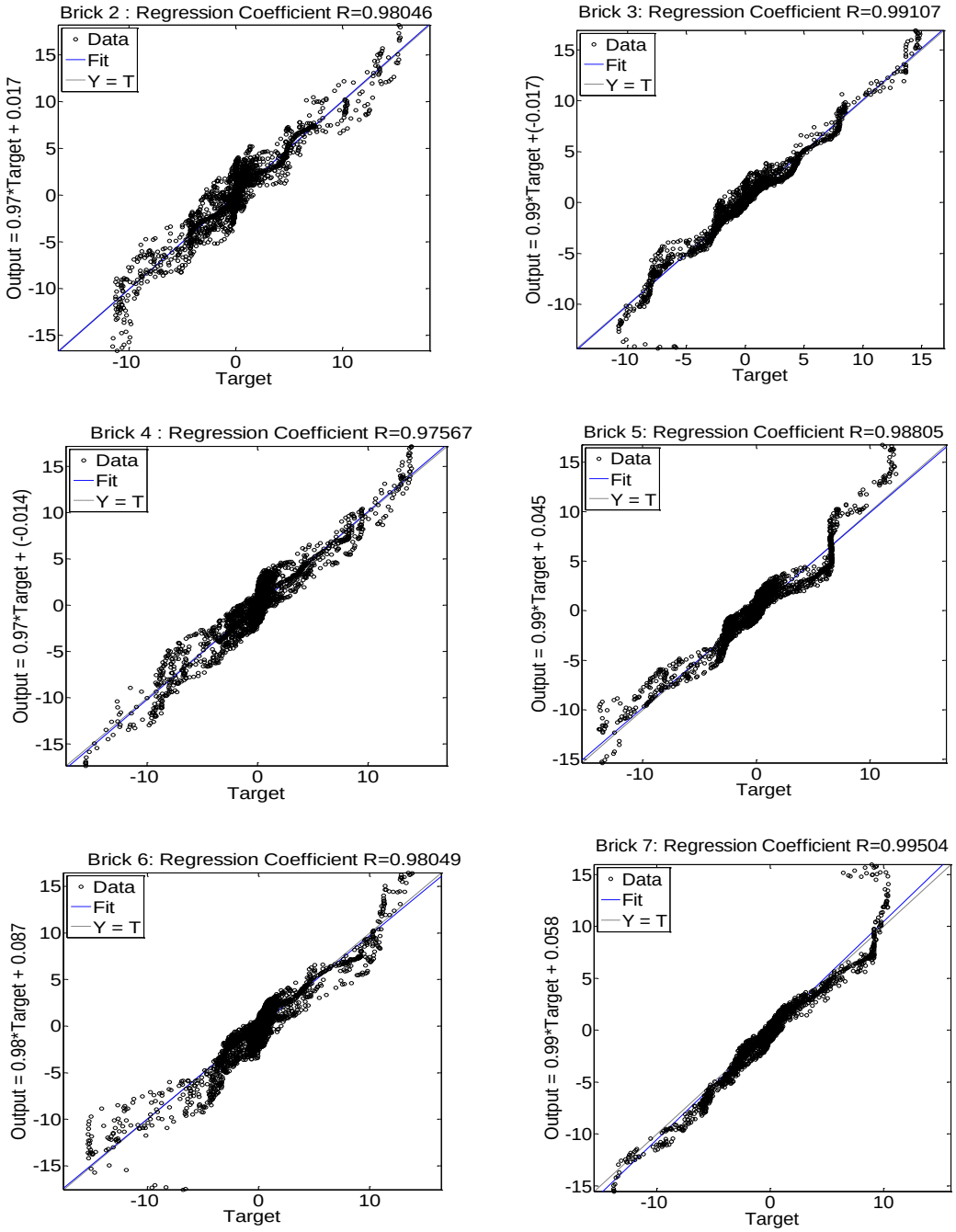


Figure 12 Regression analysis between NN prediction (Output) and experimental target data (Target) for displacement of Bricks 2, 3, 4, 5, 6 and 7 for 10 seismic tests, after 50000 epochs of training, for MLA2.

The NN predicts the displacement well for the selected locations, with all the regression coefficients ***'R'*** greater than 0.96. In Figure 13 which shows the NN performance versus brick location, the least accurate prediction (R=0.97) is for Brick 1 which is situated in the centre of the array and the most accurate predictions (R=0.99) are for Brick 3, 5 and 7 which are situated closer to the rigid array boundary. The bricks in the central part of the array have more freedom to move due to the accumulated clearance between the bricks. Their predominantly uniaxial movement is also affected by rotation and small vibrations on directions other than the direction of input application. The latter are associated with the specific design of the end face of the brick that presents a 'feature' with a prefential direction of rocking.

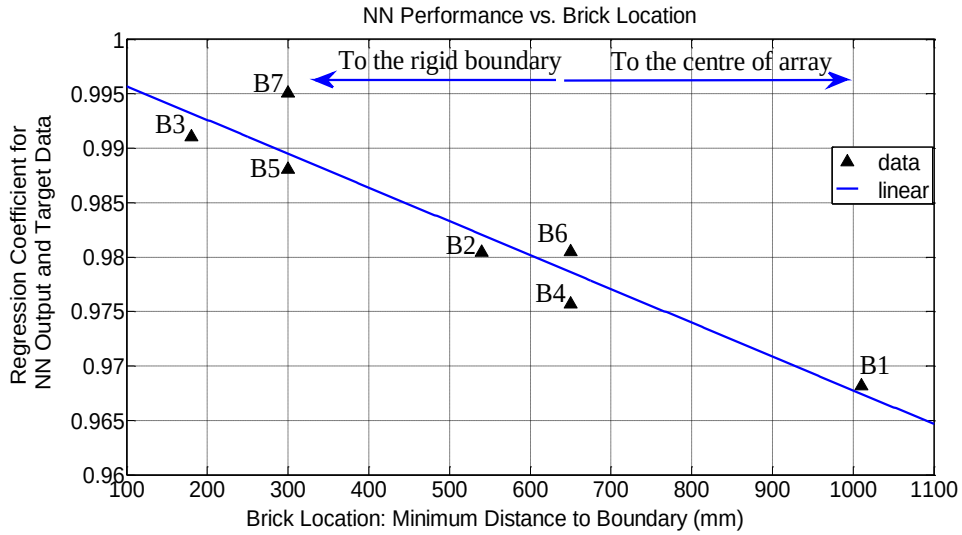
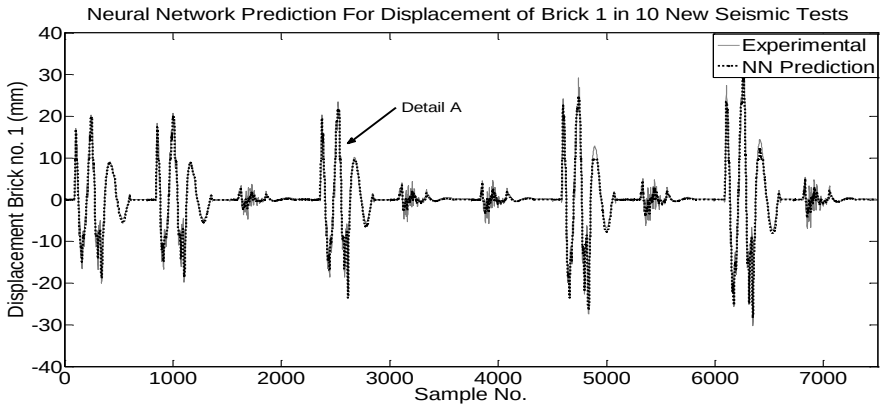
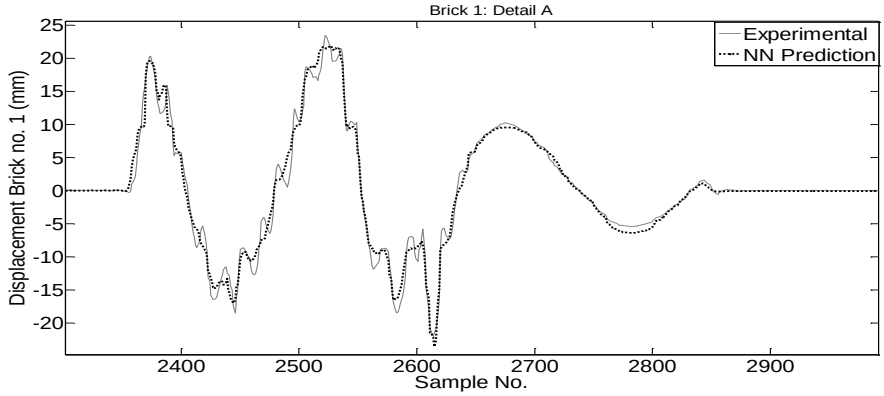


Figure 13 NN performance vs. brick location for Bricks 1, 2, 3, 4, 5, 6 and 7 (noted B1 to B7) , for 10 seismic tests, after 50000 epochs of training.

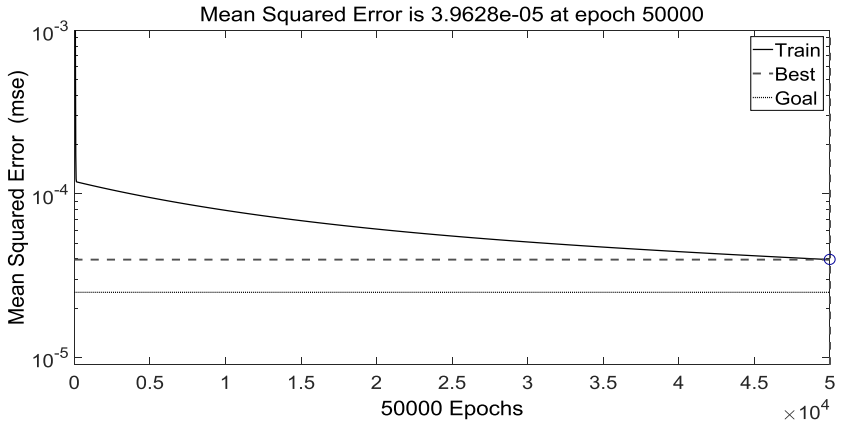
In order to test the prediction capabilities of the NN1 model, a new set of 10 seismic waves was presented to the network. The NN prediction for Brick 1 displacement compares well with the experimentally measured displacement (Figure 14). In order to assess the uncertainty of the NN prediction compared to the experimental results, the mean squared error (MSE) was calculated between the NN predicted values and the target values from the experimental set. The MSE plot (Figure 14 bottom) describes the evolution of the average of the squares of the errors as a function of the number of epochs of training. After 50000 epochs of training, the MSE reaches a value of 3.9628e-05, which was considered to be a satisfactory predicting performance for the neural network. The NN's performance appears satisfactory for this previously unseen set of inputs, increasing the confidence in the method.The NN1 model is a dynamic model, i.e. its prediction capabilities can be enriched by training when more data are fed into it. As the shaking table programme of testing progresses more, more experimental data will become available to train and increase the NN precision.



a)



b)



c)

Figure 14 NN predicted displacement for Brick 1 for a new set of 10 seismic inputs (a: general view, b: detail, c: mean squared error of prediction).

3.3.2 NN2 – Model for Relative Displacement Correlation Mapping

The NN2 architecture is similar to the architecture employed by the NN1. The NN2 input is the experimentally-measured relative displacement of Brick 1, situated close to the centre of the array. The NN2 outputs are the relative displacement of eighteen selected bricks in the top layer (B2-B19) (as marked in Figure 5).The regression coefficients between the relative displacement of Brick 1 and the other selected bricks in the top layer are a result of the NN post-processing routine. The NN2's architecture presents 3 hidden layers of 8 neurons each and an output layer of 1 neuron. Its modus

operandi is similar to the one employed by the NN1 (see Figure 7). Figure 15 presents the experimental and the NN-predicted relative displacement correlation maps for the top layer in the MLA1 (intact array), in the MLA4 (50% cracked array) and in the MLA8 (cracked brick cluster), for a sequence of 10 seismic inputs. The results are presented for tests conducted on X direction. The patterns of correlation shown in the experimental and the NN-predicted maps are similar, confirming the validity of the NN method as a powerful tool of pattern recognition and prediction. The relative displacement response of the bricks situated in the central area is better correlated to the reference brick (Brick 1) displacement, while the boundary bricks present a lower degree of coupling with the central brick. The boundary brick response is governed mainly by the input, while the bricks in the central area are also affected by a combination of geometrical effects such as brick-to-brick gap accumulation and end-face rocking features and clearances. It is important to note that, in spite of a big difference of cracking percentage between the MLA1(intact) and the MLA4 (50% cracked), the top layer appears to have an 'anchoring' effect on displacement and its response maps seem insensitive to the cracking in the layers below. The cracking in the MLA4 was randomly distributed across the array, while that in the MLA8 was randomly concentrated in a few columns close to each other: the MLA8 contained a cluster of 58 cracked bricks in the north part of the array, while the rest of the array consisted mainly of intact bricks. The correlation map between displacements show similar patters for the MLA4 and the MLA8 (Figure 15), demonstrating again that the top layer is insensitive to the level and the distribution of cracking inside the array.

Additional displacement correlation maps are shown for MLA3 (30% cracked) and MLA4 (50% cracked) in order to assess the influence of input direction on the top layer response (Figure 16). The correlation maps are given for tests conducted on X and Y direction, as shown in Figure 16. It is important to observe that the correlation maps bear the mark of the input direction, exhibiting an elongated shape in the direction of shaking. The experimental and the NN-predicted maps show similar patterns, with enhanced correlation in the central area of the array and lower correlation at the boundary. The level of cracking, again, was seen not to affect the response of the top layer. Overall, the NN2 results presented in Figures 15 and 16 increase the confidence in the NN method. All the NN predictions described in Figures 15 and 16 presented prediction errors smaller than $4e10^{-5}$ after 50000 epochs of training. As the programme of testing continues, several other array configurations will be subject of scrutiny via the NN method.

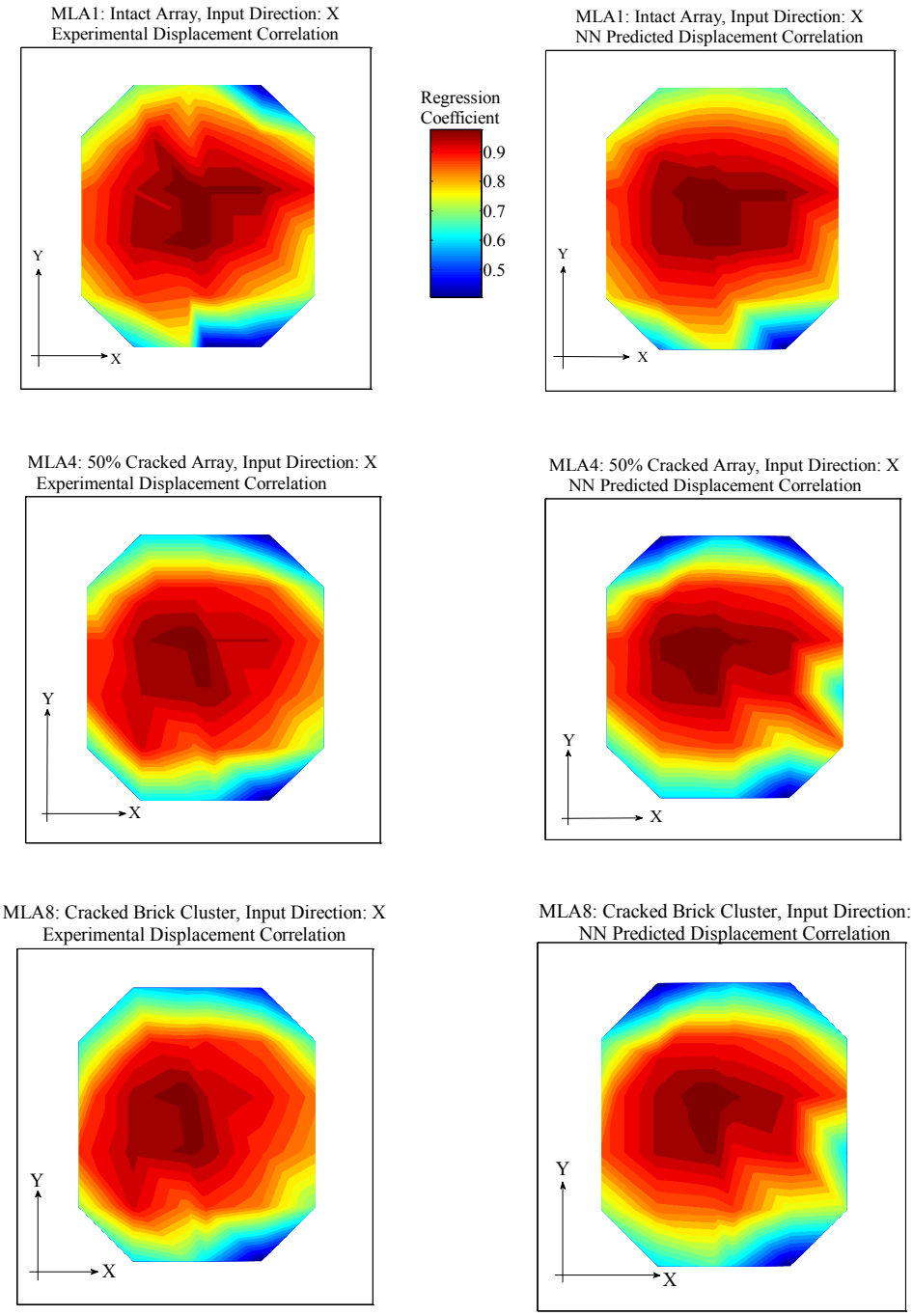
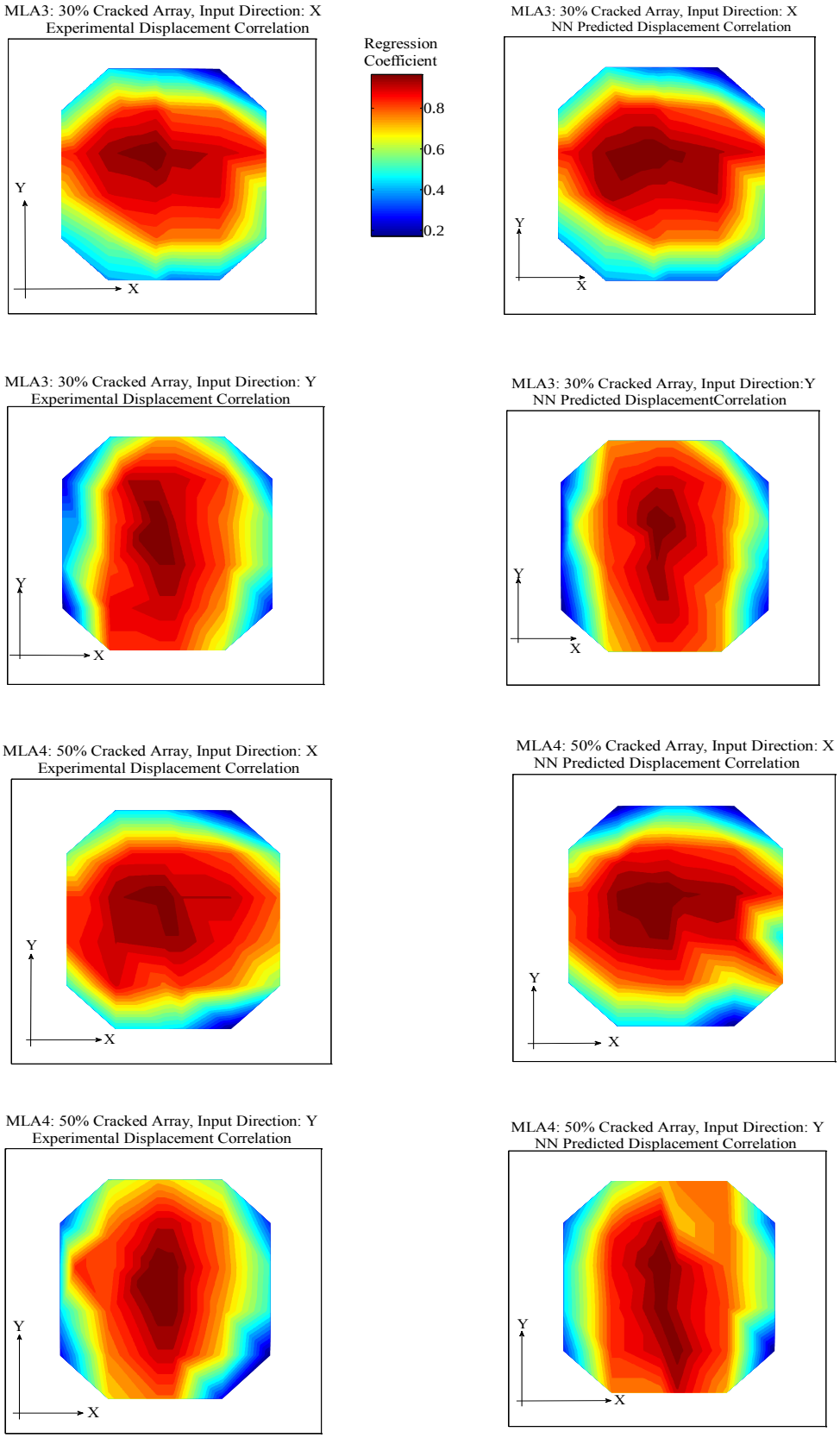


Figure 15 Correlation between relative displacement of Brick 'n' and relative displacement of Brick 1 in the top layer in MLA1 (intact), MLA4 (50% cracked) and MLA8 (5% cracked with brick cluster) for 10 seismic tests: a: experimental correlation maps . b: NN predicted correlation maps, after 20000 epochs of training. Seismic input applied on X direction.



a) b)

Figure 16 Correlation between relative displacement of Brick 'n' and relative displacement of Brick 1 in MLA3 (30% cracked) and MLA4 (50% cracked), for 10 seismic tests: a): experimental correlation maps, b):NN predicted correlation maps, after 20000 epochs of training.

3.4 Neural Network Performance

The NN performance was investigated with respect to the parameters employed in training. In general, the learning rate and the number of training iterations are problem dependent. A larger and noisier data set would require a smaller learning rate and a larger number of epochs of training (Chan et al 1995, Eaton and Olivier 1992). The above simulations employed between 1000 and 50000 training iterations. It was found that the NNs reached a stable solution after a number of epochs varying between 1000-20000 epochs of training, depending on the brick location. For example, the NN2 model dealing with relative displacement prediction is converging with more difficulty for the bricks situated in the central area than for the bricks situated near the boundary. Therefore, the maximum iteration number (20000) was used in further simulations with varying learning rate. Theoretically, if the rate at which the weights are adjusted (η) is progressively decreased during training, then the network will be able to avoid the danger of overshooting while moving its outputs closer to targets. For example, the high frequency minor oscillations that are not predicted well by the neural network (see Figure 14b) are believed to be missed because the learning rate is not sufficiently small. Within the gradient descent algorithm, a large learning rate (η) means large steps across the weight and error space. This could cause the NN to oscillate about the minimum of the error surface (Figure 17).

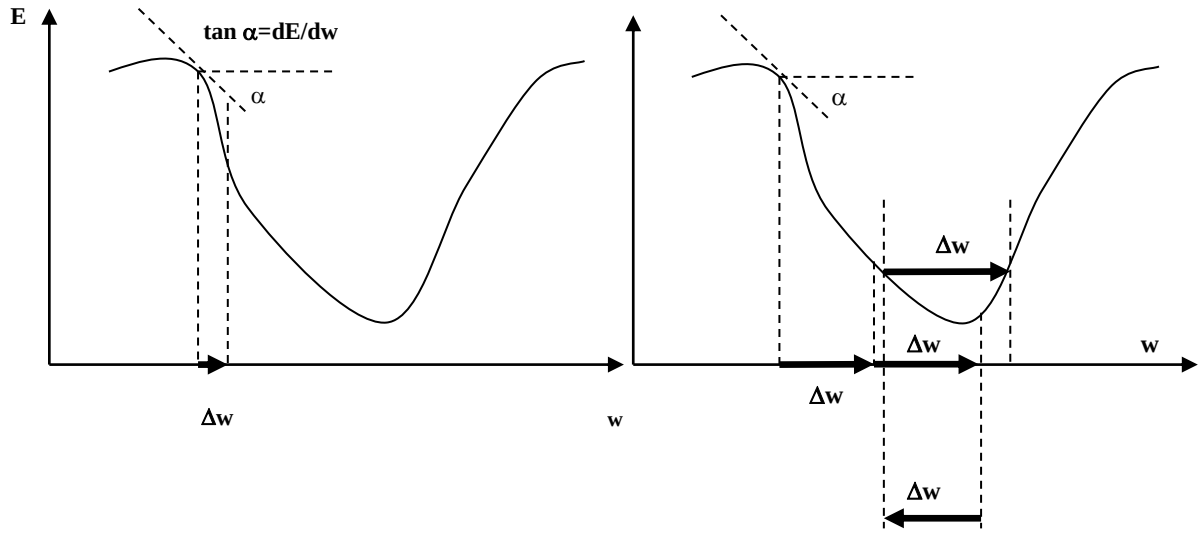


Figure 17 Oscillations about the global minimum of the error function, when the learning rate is not small enough (***E*** global error after ***k*** sessions of training, ***w*** weight between two neurons in the network, α the gradient of the error function, Δwk the increment of weight ***w*** after ***k*** epochs) (Veleenturf 1995).

When the learning rate is not suitably small there is also a danger of the system settling in a stable solution of local minima that slows down or stops the convergence process and that does not necessarily provide the correct output. When the NN gets stuck in local minima of the error surface,its performance hits a level that cannot be improved by increasing the number of iterations. In order to help the NN to find the deeper minima of the error surface, many authors recommend choosing very small learning rates to begin with (Beale & Jackson 1990 , Ellison 1992 , Miller et. Al 1990), in spite of increasing the allocated time for training. In addition to using a small learning rate, another technique to help the NN to find the global minimum of the error function involves the repetition of the learning process several times with different initializations of the weight vector. The NN is then able to determine whether the final solution is a local minimum or the final global minimum. Even with random initializations, it is possible that different initial weight vectors give almost the same solution after training. In order to avoid this situation, the authors chose initial weight vectors that were equally-spaced over the weight space. It was then straightforward to guarantee that the

subsequent initializations were different. Figure 18 shows an example of the initial distribution of the weight vector associated with the first input and the eight neurons in the first layer.

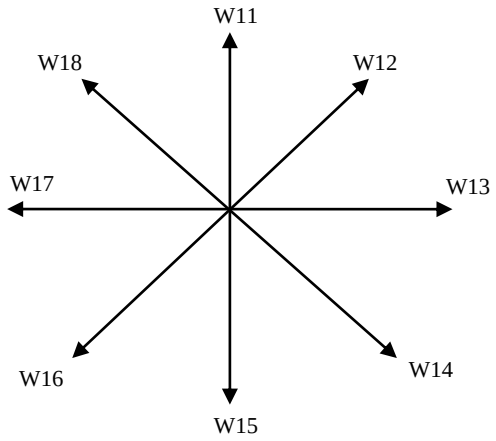
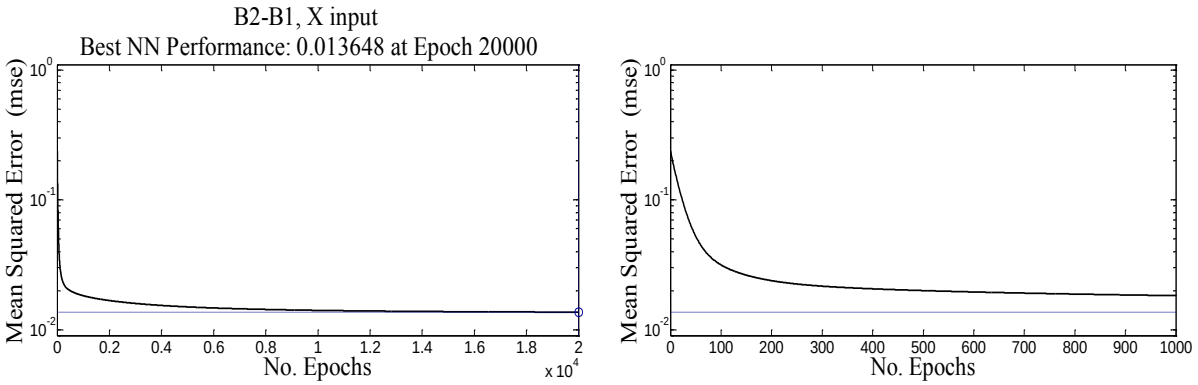


Figure 18 Example of the initial distribution of the weight vector 'w1i' (i=1:8).

When the learning process was repeated for the second time, the set of initial weight vectors were turned through a specified angle (e.g. 30^o). The same procedure was applied for the subsequent initializations. In this manner, the NN was able to find the deepest minimum of the error surface. Figure 19 shows the performance function of the NN2 model predicting the correlation between the relative displacement of Brick 2 and Brick 1. It appears that the performance function is almost reaching a plateau after only 1000 epochs (Figure 19b). However, the continuation of training is necessary up to 20000 epochs, when the performance function is reaching its goal(Figure 19a).



a) b)

Figure 19 Error function during training. Correlation of relative displacement for Brick 2 and Brick 1, in MLA3 array, for 10 inputs on X direction. a: best NN performance, b: detail- performance function after 1000 epochs.

The present results are considered encouraging. Future work may involve an exploration into how a change of the training algorithm can enhance the prediction performance of the models.

4. Concluding Remarks

This paper presents an NN approach for displacement analysis in a complex multi-body assembly subjected to seismic loading. Component displacement has a crucial impact on the local distortion and on the channel shapes of the AGR model. The NNs can predict component displacements and

displacement correlation maps for user-defined dynamic inputs and the predictions compare well with the experimentally obtained parameters. Such predictions are very important in understanding the interaction pattern within the MLA array, as the response can be mapped for new seismic inputs.
The advantages of the NN models are:

a) NNs do not need any given assumption regarding the relation between inputs and outputs and are capable to recognize patterns and generalize. They can be trained with a large number of seismic waves to test the response for a wider case.
b) NNs could play a role in the design of experimental programmes: if the NN predictions reveal peculiarities for certain seismic inputs, then those inputs could be employed in the physical experiment to measure the response.
c) NNs are general purpose in nature, being capable to adapt to a wide variety of test and array configuration scenarios.
d) NNs have the ability to train from the given tests, even if data may be sparse, incomplete or noisy.

The limitations of the NN models are:

a) A large amount of experimental data is required, which takes long to assemble. The NN training is time consuming and architecture dependent.
b) NNs lack explanation facilities and behave as 'blind' systems during training.
c) NNs need a wide enough range of training scenarios, as insufficiently rich datasets may lead to wrong predictions and decisions.

This paper presents a first foray into modelling the MLA data via neural networks. The authors take heart from this first set of NN models which can be dynamically trained to accumulate more data and enhance their precision, as the shaking table programme of testing continues. A similar approach will be explored for predicting the channel shapes under seismic loading and a combination of neural networks with other associative memory techniques may also be pursued for experimental design and 'big data' investigation.

Acknowledgement

The authors would like to thank EDF Energy for financial and technical support. The views expressed in this paper are those of the authors and do not necessarily represent those of EDF Energy.

List of Abbreviations

AGR: Advanced Gas-cooled Reactor
IR: Infrared
MLA: Multi Layer Array
MSE: Mean Squared Error
NN: Neural Network
NNC: National Nuclear Corporation
RRS: Required Response Spectrum
SD: Standard Deviation
UOB: University of Bristol

References

Bartal Y, J. Lin J., Uhrig R. E., (1995), *Nuclear power plant transient diagnostics using artificial neural networks that allow "don't know" classifications,* Nuclear Technology, Vol. 110.

Beale R. and Jackson T.(1990), *Neural computing: an introduction*, Bristol and Philadelphia: Institute of Physics Publishing.ISBN 0852742622 9780852742624.

Beale M.H., Hagan, M.T., Demuth H.B., (2016), *Neural Network Toolbox User's Guide. R2016b*, © Copyright 1992–2016 by The MathWorks, Inc.

Bourguet R.E. and Antsaklis P.J., (1994), *Artificial Neural Networks In Electric Power Industry*, Technical Report of the ISIS Group at the University of Notre Dame ISIS-94-007 April, 1994.

Chan W.T., Chow Y.K. and Liu L.F. (1995) *Neural Network: An Alternative to Pile Driving Formulas*. Computers and Geotechnics, 17(4):135-156.

Cheon, S.W., S.H. Chang S. H, (1993), *Application of neural networks to a connectionist expert system for transient identification in nuclear power plants*, Nuclear Technology, Vol. 102.

Cybula (2016): http://cybula.squarespace.com/contact/

Dihoru L, Crewe, A.J, Taylor, C.A., Horgan, T. (2011), *Shaking Table Experimental Programme*, Proc. of the Conference on Modelling and Measuring Reactor Core Graphite Properties and Performance, Aston University,. Edited by Gareth B. Neighbour, Cambridge, UK.

Dihoru L., Oddbjornsson O., Crewe A.J., Taylor C.A., Horgan T., Webster E., (2012), *Validation of Graphite Core Numerical Models*, Proc. of the 13th International Nuclear Graphite Specialist Meeting, Meitingen, Germany, 23-26 Sept.

Dihoru L., Oddbjornsson O., Horseman T., Dietz M., Wilson J., Kloukinas P., Voyagaki E, Crewe A.J., Taylor C.A. (2014), *Multi-Layer Array Rig Work For Seismic Behaviour With Cracked Bricks*, Proc. of the 4th EDF Energy Nuclear Graphite Symposium – Engineering Challenges Associated With The Life Of Graphite Reactor Cores, Nottingham, Edited by Flewitt P.E.J. and Wickham A.J., EMAS Publishing.

Dihoru L., Oddbjornsson O., Kloukinas P., Dietz M., Horseman T., Dietz M., Voyagaki E, Crewe A.J., Taylor C.A. (2016), *The development of a physical model of an Advanced Gas Cooled Reactor core: Outline of the feasibility study*, Nuclear Engineering and Design, 323 (2017) 269–279.

Eaton H.A.C. and Olivier T.L. (1992) *Learning Coefficient Dependence on Training Set Size*, Neural Networks, Vol. 5, p. 283-288.

Ellison, D., (1992) *'Generalisation and convergence in the multi-dimensional Albus Perceptron (CMAC)'*, Proceedings of the Second Irish Neural Networks Conference, Northern-Ireland, 25-26 June 1992, p.199-208.

Flewitt and Wickham (2015) (eds), *Engineering Challenges Associated with the Life of Graphite Reactor Cores*, EMAS Publishing, ISBN-978-0-9576730-5-2.

Hagan M.T. , Demuth H.B., Beale M.H., De Jesus O, (2014), *Neural Networks Design (2nd Edition)*, Publisher: Martin Hagan, ISBN-10: 0971732116

Heaton, J.(2012), *Introduction to the Math of Neural Networks*, Publisher: Heaton Research, Inc., ASIN: B00845UQL6

Hornik, K., Stinchcombe, M., and White, H.,(1989), *Multilayer Feedforward Networks Are Universal Approximators*, Neural Networks, Vol. 2, p.359.

Koziara T and Bićanić N (2011), *International Journal for Numerical Methods in Engineering*,Vol. 87, p. 437–456.

Kralj, B., Humphreys, S.J. and Duncan, B.G.J. (2005). *Proc. of the Conference on Seismic Modelling of an AGR Nuclear Reactor Core*, Cardiff, United Kingdom, p. 193-200.

Miller, W.T., Sutton, R.S. and Werbos, P.J.,(1990), '*Neural Networks for Control'*, The MIT Press, Cambridge, Massachusetts.

Neighbour, G.B. (2007) (ed), *Management of Aging Processes in Graphite Reactor Cores*, RSC Publishing, Cambridge. ISBN-978-0-85404-345-3.

Neighbour, G.B. (2013) (ed), *Modelling and Measuring Reactor Core Graphite Properties and Performance*, RSC Publishing, Cambridge, ISBN-978-1-84973-390-8.

Ohga Y., Seki H., (1993), *Abnormal event identification in nuclear power plants using a neural network and knowledge processing*, Nucl. Technol, Vol. 101.

Principia Mechanica Limited, (1981), *Seismic Ground Motion For UK Design (Report)*,archive of British Nuclear Fuels Limited and Central Electricity Generating Board, April 1981.

Reifman, J., Wei, T. Y. C., Abboud, R. G., and Chasenky, T. M., (1994), *Cooperative Research and Development for Artificial Intelligence ReactorDiagnostic System*, Proc. Am. Power Conf., Chicago, Illinois, April 25-27, 1994, Vol. 56-I, p.365, Illinois, Institute of Technology, Chicago, Illinois.

Reifman, J., Wei, T. Y. C., Vitela J.E., Applequist C.A., and Chasenky, T. M., (1996), *PRODIAG - A Hybrid Artificial Intelligence Based Reactor Diagnostic System for Process Faults,* Proc. Fourth Int. Conf. Nuclear Eng., New Orleans, Louisiana, March 10-14, 1996, Vol. 1-B, p. 885, Rao A.S., Duffey R.B., and Elias D. , Eds., The American Society of Mechanical Engineers, Salem, Massachusetts..

Santosh T.V. Vinod, G, Saraf R.K, Ghosh A.K., Kushwaha H.S., (2007), *Application of artificial neural networks to nuclear power plant transient diagnosi*s, Reliability Engineering & System Safety, Vol. 92, Issue 10, p. 1468–1472

Uhrig, R. E., (1991a), *Potential Application of Neural Networks to the Operation of Nuclear Power Plants*, Nucl. Safety, Vol. 32, Issue 68.

Uhrig, R. E., (1991b), *Neural Networks and Their Potential Application to the Operation of Nuclear Power Plants*, Proc. Topical Meeting Frontiers in Innovative Computing for the Nuclear Industry, Jackson Lake, Wyoming, September 15-18, 1991, Vol.II, p. 439, American Nuclear Society, La Grange Park, Illinois.

Veelenturf L.P.J., *Analysis and Applications of Artificial Neural Networks*, (1995), Prentice Hall Intl (UK) Ltd., p.70-76

Westwick D (ed) (2007), *Power Plants and Power Systems Control 2006:Proceedings of the IFAC Symposium on Power Plants and Power Systems,* Kananaskis, Canada, ASIN: B008KNZ8B0.